DOCUMENT RESUME

ED 342 727 SP 033 516

AUTHOR Boone, William J.
TITLE How Psychometric Analysis Can Help Teachers Realize a Curriculum.
PUB DATE 90
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Curriculum Development; Elementary Education; Elementary School Science; *Item Analysis; Item Response Theory; *Psychometrics; Relevance (Education); Student Evaluation; *Test Interpretation; *Test Use
IDENTIFIERS *Curriculum Mapping; *Rasch Model

ABSTRACT

Classroom teachers use intuitions, experiences, syllabi, and course guides to help implement a curriculum. Psychometric techniques can also help guide and shape curricula. Some psychometric techniques allow tests to reveal the actual curriculum mastered by students in an objective and sample-free manner and to indicate a probable learning sequence of topics. The graphical presentation of such data is called a "curriculum map." The power of the maps is such that by displaying students' mastery of material and the difficulty of specific topics, teachers may be better guided in an optimal presentation order of material to students. Curriculum maps are built by calculating item difficulties and student abilities. The calculations are important because the nonlinearity of raw test scores are corrected to make the data from students and items meaningful. The paper introduces a number of psychometric curriculum maps: (1) real item maps; (2) student and item maps; and (3) fit maps of students and items. All of the maps can be constructed easily with modern psychometric techniques. They are a logical, necessary step in evaluating curriculum, and they provide the opportunity for educators to improve and study topic sequencing. (SM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED342727

HOW PSYCHOMETRIC ANALYSIS CAN HELP TEACHERS REALIZE A CURRICULUM

WILLIAM J. BOONE
SCHOOL OF EDUCATION
INDIANA UNIVERSITY-BLOOMINGTON

PAPER PRESENTED AT THE 1990 NATIONAL AERA MEETING

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W. Boone

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SP033 516

BEST COPY AVAILABLE

William J. Boone
Program in Science Education
School of Education 204-C
Indiana University
3d and Jordan Ave.
Bloomington, IN 47405
(812) 855-8658

HOW PSYCHOMETRIC ANALYSIS CAN HELP TEACHERS REALIZE A CURRICULUM

Anyone who has taught or observed an elementary school teacher realizes the enormity of the curricular task confronting instructors at this level. In addition to dealing with the emotional and psychological needs of the students, each teacher is expected to have mastered and to present lessons on a wide variety of subjects (Mathematics, Science, Social Studies). Unfortunately, as elementary teachers prepare in college it is impossible to complete all of the academic courses which would provide the theoretical and practical experience for each elementary field. This reality makes the development of useful and varied curriculum guides exceedingly important.

Instructors who must plan, present, and access a range of subject material once in the classroom are presently aided by various tools which help them implement a curriculum. One device teachers use, from the minute they enter the classroom, is a "gut" feeling from interactions and assessments of students. By using their intuitions,

instructors gauge how difficult or easy specific topics "really" are for students, and then consequently use this information to pace and direct a curriculum. Two other tools commonly utilized by teachers to orient their teaching are syllabi (from school and district) and publishers teaching guides. Syllabi outline material which must be covered, and thereby provide a checklist for the teacher. Teaching guides commonly include tips to instructors on the way in which material should be presented in classes, and warn teachers of student pitfalls. Almost uniformly, publishers suggest that the chapters of a text be presented in numerical order. Many teachers, especially those with a limited foundation in the topics which must be presented, use the publisher's suggested curriculum to the letter.

Teachers in the classroom have their intuitions, experiences, syllabi, and course guides to help in the implementation of a curriculum, however, psychometric techniques can be used to guide and shape curricula. Some psychometric techniques allow tests to reveal 1) the actual curriculum mastered by students in an objective and sample-free manner and 2) a probable learning sequence of topics. The graphical presentation of these data will be called "curriculum maps". The power of these "maps" is that by displaying students' mastery of material, and the difficulty of specific topics, teachers may be better guided in an optimal presentation order of

material to students. By using such techniques a key curriculum problem noted by P.H. Phenix can begin to be solved. He wrote "The problem of curriculum is to economize scarce learning potential by making the most judicious and appropriate selection of study content. Human intelligence is too rare and precious a thing to squander on a haphazard program of instruction".

Theoretical Framework Behind Curriculum Maps

The "curriculum maps" which will be presented are built by calculating "item difficulties" and "student abilities". These calculations are important because the non-linearity of raw test scores must be corrected so as to make the information from students and items meaningful. The non-linearity of raw scores can be understood by noting that a 5 point improvement of a student's score from a 90 to 95 on a 100 item test represents three times the improvement as that indicated by another student's 5 point gain from 50 to a 55 on the same test. By correcting for non-linearity a metric picture (a map) of the achieved curriculum can emerge. Evaluation of data with these techniques can be, within limits, sample and item independent.

The theory and implementation of Rasch Measurement (the techniques which have been used to derive these initial curriculum maps) has been developed (Rasch, 1960/1980; Wright and Stone, 1979; Wright and Masters, 1982) and applied successfully to a wide variety of data (Knox Cube Test of Short Term Memory, Wright and Stone, 1983; The National Reference Scale for Reading: An Application of the Rasch Model, Rentz and Bashaw, 1977; Item Banks: What, Why and How, Wright and Bell, 1984; Setting, Evaluating, and Maintaining Certification Standards with the Rasch Model, Grosse and Wright, 1986).

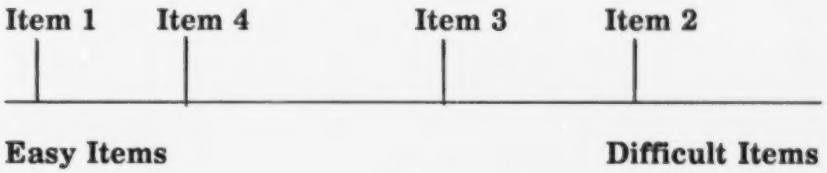
Curriculum Maps

What are the generic considerations by which a curriculum might be sequenced? What does the difficulty of test items for students tell about the way in which topics should be presented in class? When are students required to make conceptual leaps in understanding? All of these questions can be attacked by using psychometric techniques to map the "achieved" curriculum and thereby specify and compare the relative difficulties of test items which have been calibrated from student responses.

In the remaining part of this paper a number of psychometric curriculum maps will be introduced and curriculum implications discussed. It is important to keep in mind that all of the maps discussed 1) can be easily constructed with modern psychometric techniques; 2) are a logical and necessary step in the evaluation of curriculum- particularly with respect to the relation between what is intended (in texts and by teachers) and what is actually achieved; 3) provide a unique opportunity for educators to improve and study topic sequencing.

Types of Curriculum Maps

The psychometric technique (Rasch Analysis) used to build curriculum maps converts non-linear "raw scores" to measures on a linear logit scale. One way of presenting the results of an analysis is to map items onto a unidimensional scale which represents the variable of a test. A schematic item map illustrating this idea is presented below:



Item maps such as the one presented on the preceding page furnish important curricular advice in two key ways. First, the ordering of items from easy to difficult suggests a natural ordering of material in class. Knowing what areas of a curriculum have been mastered by students, and what topics need additional work would enable realistic and achievable goals for teachers and administrators to be reached. It is true that text book publishers try to present material from easy to hard, and thus furnish teachers with a great deal of curricular guidance, but teachers need a tool, such as these item maps, which displays the empirical difficulty of topics for students. Administrators too could be greatly aided by item maps when the feasibility and sequencing of school and district teaching goals are reviewed and revised.

The ordering of items from "easy" to "difficult" is not the only piece of valuable information these maps contain. The size of the gaps between items on the variable line are also of great assistance to any curriculum planner, for each gap represents a jump in difficulty from topic to topic. Explicit information on the points in a curriculum leaps of understanding are required of students is valuable for both veteran and novice teachers.

Real Item Maps

A curriculum map is presented in figure 1. This map was built from a set of 90 science items taken by over 800 4th graders in the fall of 1988 and by a similar number of 5th graders in the fall of 1989. The items used in this map are presented in table 1. With this map much can be learned about how a science curriculum might be structured. What sort of information is highlighted with this curriculum map?

The positions of the "unit" questions are revealing. In most texts and teaching guides "volume" is presented after "distance". Sometimes the concept of volume is presented after "mass". What does this map tell us? Simply that volume could be presented as one of the first measurement concepts to students at the 4th and 5th grade level. Volume may be "easy" because students have everyday experiences with volume through milk cartons and 2 liter Pepsi bottles. The curriculum moral? Present familiar measurement topics first, and recognize that volume may not be so hard for students! The calibration of items 52-54 and 7-9 illustrate it is fairly easy for 4th and 5th graders to identify properties of "things" they are familiar with. Properties might range from how something looks, to how something changes (grows).

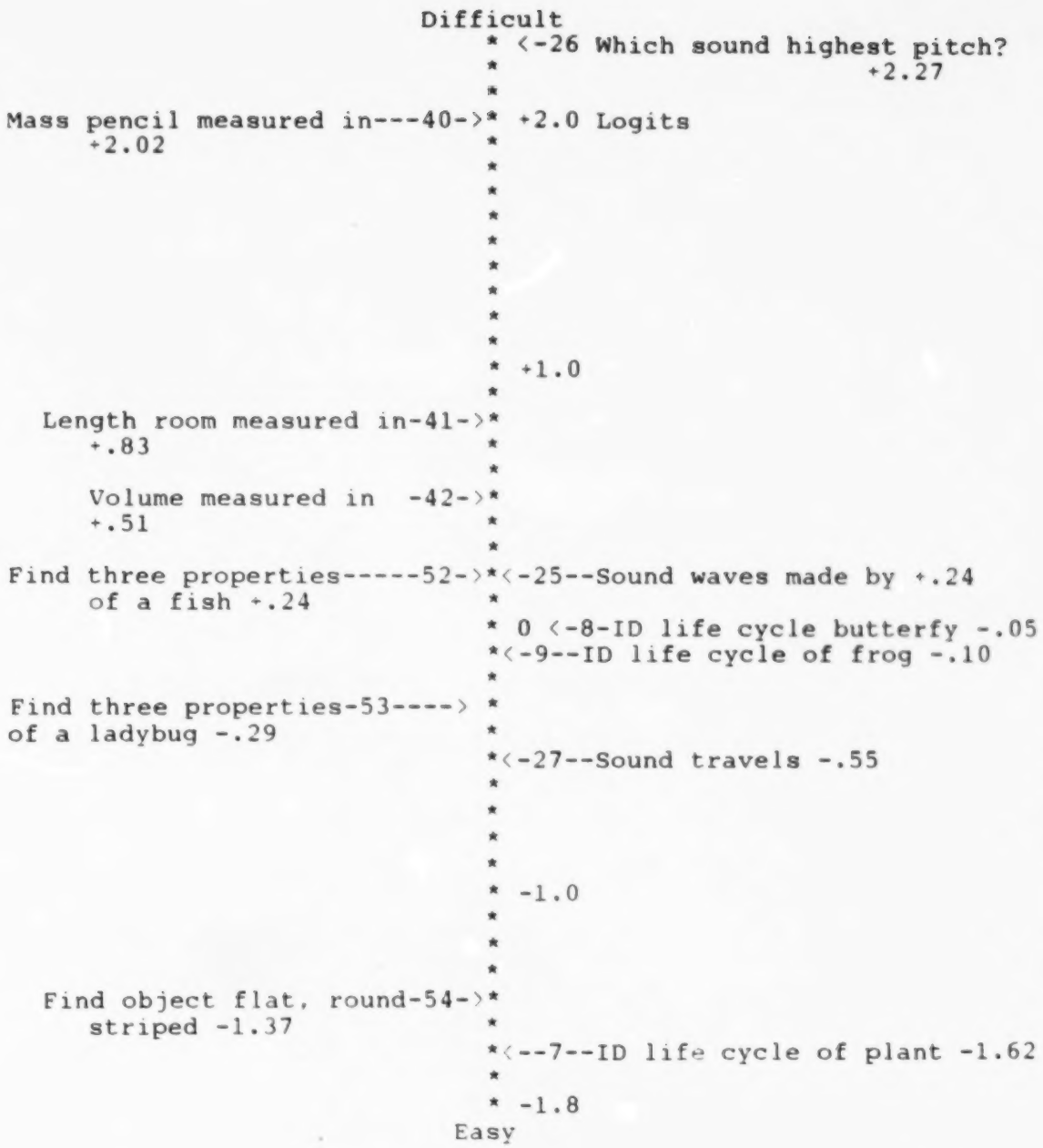


Figure 1. Sample curriculum map.

Table 1. Text of Selected Test Items Used to Construct a Sample Curriculum Map

---

Item 7:
Put the pictures in order to show the life cycle of a bean plant.
(Four numbered pictures of a bean plant at different stages of growth shown.)
Response Options-
A 4, 2, 3, 1
B 2, 3, 4, 1
C 4, 2, 1, 3
D 3, 1, 2, 4

Item 8:
Put the pictures in order to show the life cycle of a butterfly.
(A picture of a butterfly, larvae a leaf, cocoon, and caterpillar are shown.)
Response Options-
A 2, 3, 1, 4
B 2, 4, 3, 1
C 2, 1, 4, 3
D 2, 3, 4, 1

Item 9:
A tadpole hatches from an egg, grows into a frog, and dies. These changes are part of its
Response Options-
A water cycle
B life cycle
C food chain
D food web

Item 25:
Sound waves are made by
Response Options-
A heat.
B light.
C gravity.
D vibrations.

Table 1 (Continued). Text of Selected Test Items Used to Construct a Sample Curriculum Map

Item 26:
Which sound has the highest pitch?
Response Options-
A thunder
B a fog horn
C a whistle
D a bass drum

Item 27:
Sound travels
Response Options-
A sideways only.
B up and down only.
C in all directions.
D in one direction only.

Item 40:
The mass of a pencil can be measured in
Response Options-
A grams.
B meters.
C liters.
D Celsius.

Item 41:
The length of a room is measured in
Response Options-
A grams.
B meters.
C liters.
D Celsius.

Table 1 (Continued). Text of Selected Test Items Used to Construct a Sample Curriculum Map

---

Item 42:
The volume of a milk carton can be measured in
Response Options-
A grams.
B meters.
C liters.
D Celsius.

Item 52:
Find three properties of this fish.
(A black and white line drawing of a fish is provided.)
Response Options-
A soft, curvy, spotted
B rough, red, hard
C soft, slippery, square
D soft, curvy, striped

Item 53:
Find three properties of this ladybug.
(A black and white line drawing of a ladybug is provided.)
Response Options-
A spotted, round, black head
B spotted, square, black body
C spotted, smooth, has four legs
D round, striped, has six legs

Table 1 (Continued). Text of Selected Test Items Used to Construct a Sample Curriculum Map

---

Item 54:
Find the object that is flat, round, and striped.
(Line drawings of a rock, a square wooden block, a circular disc, and a baseball are provided. Each object was given a unique letter for identification.
Response Options-
A
B
C
D

The positions of items 7-9 are also interesting from another curricular standpoint. Many 4th grade books use butterflies to illustrate concepts. What is clear from the relative positions of this set of three problems is the tricky conceptual jump required to understand butterfly evolution. One interpretation of the three problems relative spacing is students recognize (and accept) the physical growth of a butterfly as being quite different from that of a plant. Certainly it would be important for teachers to be alerted to this learning jump.

The final set of test items used to build this map are three questions (25-27) all involving sound. One revealing aspect of these items is the ease of 27 and the difficulty of 26. Familiarity may again play a role. Students can envision (and hear) sound moving in all

directions, making 27 easy. Problem 26 may be difficult from a lack of familiarity (vocabulary) with the word pitch. How do all these items work together to define a science curriculum? One way is by showing the importance of starting a topic presentation with familiar objects and actions.

Student and Item Maps

Maps composed of items positioned by difficulty are not the only useful curriculum aid to be generated from this type of analysis. Maps that position student "measures" among item "difficulties" on a common linear scale can also be calculated (figure 2). Maps containing both calibrated items and student measures can help teachers in a most comprehensive fashion. Teachers can learn where each student's ability falls with respect to the difficulty of test items. The concepts students have and have not yet mastered can be easily seen. Student/item maps furnish a teacher with a conceptual picture of what is actually being taught and where students fall in a curriculum defined by banks of test items.

Student/item maps are important curricular tools in another way. From a curricular planning standpoint, it is of great importance to improve the quality and range of data collected from students. Because student responses to test items are often used to evaluate a

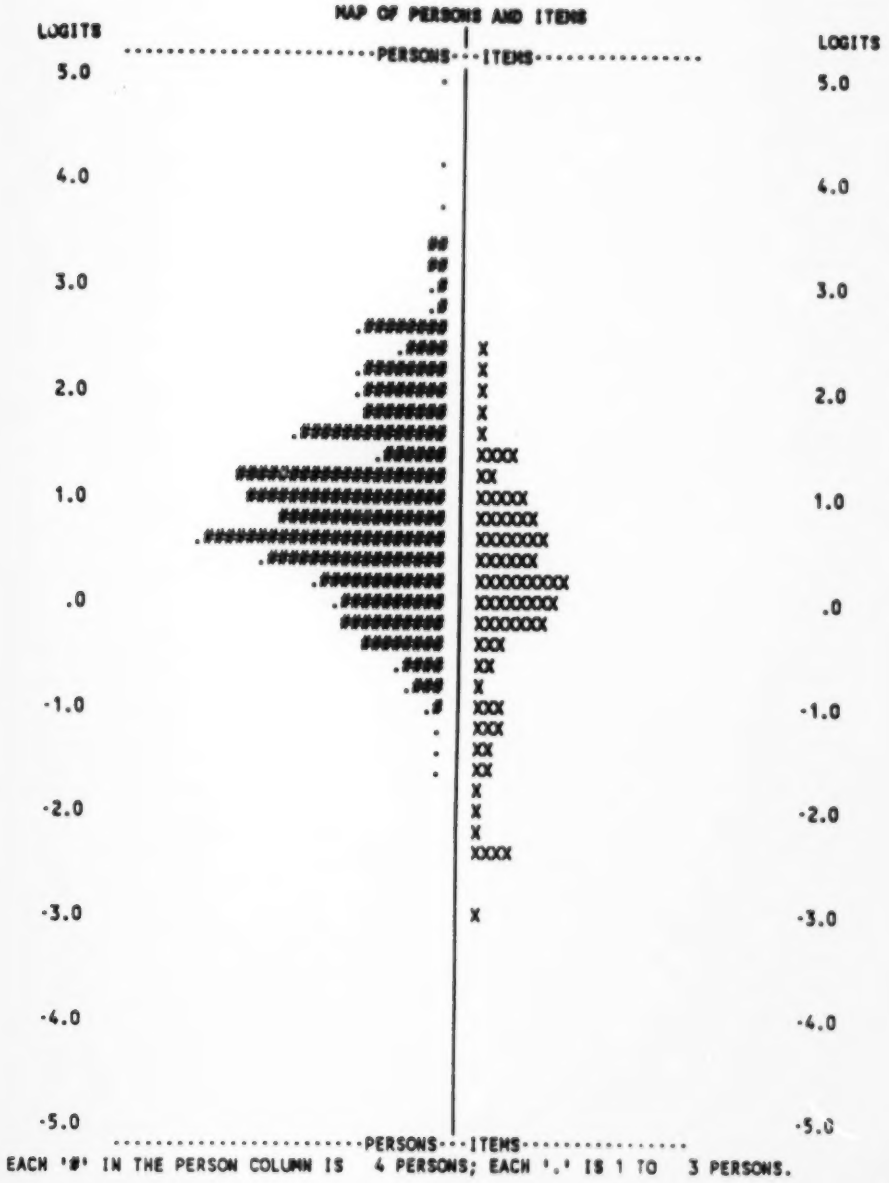


Figure 2. Sample student/item map.

curriculum, student/item maps provide critical information on the adequacy of test items used to collect student data. In the example student/item map one can clearly note an overabundance of questions in some ranges of "item difficulty", and a lack of items at other difficulty levels. If the achieved curriculum is to be gauged accurately tests must be constructed with a range of items sufficient to place a student in a curriculum. It is important to note that although this psychometric technology would most often be used to make an objective comparison of the achieved and the predicted curriculum of students from one point in time, new test data can be easily integrated into old data and new student/item maps generated. With these new maps the progress students make can be assessed and the ordering of the curriculum altered to take student learning into account.

"Fit" Maps of Students and Items

Two types of statistics (called infit and outfit) provide important information on 1) students' individual reactions to specific test items, and 2) each student's answers to a set of items. To learn more about the successes and failures of a curriculum, infit and outfit data can and must be used.

Item Outfit Maps

One important way in which teachers try to evaluate a curriculum is by spotting instances of students acting in an "unexpected" manner to an item. Detecting situations which cause a good student to get an easy item wrong, or a less able student to get a hard item correct, can facilitate the design of better tests and curriculum. Rasch analysis allows the calculation of "item outfit" statistics which indicate how inconsistently an item was used by students. To illustrate "outfit" consider some items from this 4th/5th grade science test.

Item 11 involved the question: Where do trees get their energy directly from? This item was identified as having a high outfit value (some students responded in an unpredictable way to this question). Now that the Rasch analysis has called attention to this item, the reasons for student confusion must be considered. Probably the difficulty is related to the responses offered for this item. The options for this item were a) water, b) sun, c) air and d) soil. The high outfit probably resulted from confusion students had with these options because there is <u>not</u> one right answer to this item. In this case, the outfit of the item suggests the need for improved response options. Furthermore, the outfit also calls attention to the need for careful introduction of a concept (objects "getting energy" from other

substances). Students and teachers alike may be confused as to the differences and similarities of energy and nutrients.

Another question on the test which had a high outfit value (and thus suggested student confusion) was the item: All living things have: a) heads, b) sisters, c) feet and d) parents. One explanation for the high outfit may be that students reviewed just the first response and then tried to decide whether a response was valid based on their own experiences. Students may have thought along the lines of "My gerbil is living and has a head" so answer "a" must be right. For this item students may be reading "All living things have" as "Living things have". Another possibility for this item's outfit might be students either did not understand, or had not gotten to the segment of their science curriculum involving reproduction.

Student Outfit Maps

An integral part of improving a curriculum not only involves evaluating the reaction of students to particular items, but also investigating the reaction of specific students to banks of items. When teachers grade a test, they access and revise curriculum in light of techniques that seem to have allowed poor students to understand important topics. Student "outfit" maps enable a curriculum planner 1) to identify test takers who responded strangely

and 2) to follow up by investigating the particular responses of specific students. Student outfit maps for a district may help indicate how well particular teachers present specific topics.

There are many instructional tools which can be used by teachers in the classroom as well as those in training. A psychometric analysis of these elementary school science test data suggests that objective measurements of test items and students can be used to provide important curricular advice to teachers which augments the information provided in text books and teaching guides.

References

Grosse, M.E. and Wright, Benjamin D., 1985. Validity and Reliability of True-False Tests, Educational and Psychological Measurement, V45, pp. 1-13.

McKnight et al., 1987. The Underachieving Curriculum, Champaign, Illinois, Stips Publishing Company.

Phenix, P.H., 1958. Philosophy of Education, New York, Holt, Rinehart and Winston.

Rasch, George, 1960. Probabilistic Models for Some Intelligence and Attainment Test, Copenhagen, Denmark: Danmarks Paedogogiske Institut (reprinted, University of Chicago Press, 1980).

Rentz, R. Robert and Bashaw, W.L., 1977. The National Reference Scale for Reading: An Application of the Rasch Model, Journal of Educational Measurement, V14, pp.161-179.

Stone, Mark H. and Wright, Benjamin D, 1983. Measuring Attending Behavior and Short-Term Memory with Knox's Cube Test, Educational and Psychological Measurement, V43, pp. 803-814.

Wright, Benjamin D. and Stone, Mark H., 1979. Best Test Design, Chicago, MESA Press.

Wright, Benjamin D. and Masters, Geofferey N., 1982. Rating Scale Analysis, Chicago, MESA Press.

Wright, Benjamin D. and Bell, Susan R., 1984. Item Banks: What, Why, How, Journal of Educational Measurement, V21, #4, pp.331-345.

aera.5